

treatment of other eating disorders remains to be studied.

- FAIRBURN, C. G., AGRAS, W. S. & WILSON, G. T. (1992) The research on the treatment of bulimia nervosa: practical and theoretical implications. In *The Biology of Feast and Famine: Relevance to Eating Disorders* (eds G. H. Anderson & S. H. Kennedy), pp. 317–340. New York: Academic Press.
- , JONES, R., PEVELER, R. C., *et al.* (1993a) Psychotherapy and bulimia nervosa: the longer-term effects of interpersonal psychotherapy, behaviour therapy and cognitive behaviour therapy. *Archives of General Psychiatry*, **50**, 419–428.
- , MARCUS, M. D. & WILSON, G. T. (1993b) Cognitive behaviour therapy for binge eating and bulimia nervosa: a treatment manual. In *Binge Eating: Nature, Assessment and Treatment* (eds C. G. Fairburn & G. T. Wilson), pp. 361–404. New York: Guilford Press.
- GARNER, D. M. & BEMIS, K. M. (1985) Cognitive therapy for anorexia nervosa. In *Handbook of Psychotherapy for Anorexia Nervosa and Bulimia* (eds D. M. Garner & P. E. Garfinkel), pp. 107–146. New York: Guilford Press.
- WILSON, G. T. & FAIRBURN, C. G. (1993) Cognitive treatments for eating disorders. *Journal of Consulting and Clinical Psychology*, **61**, 261–269.

CHRISTOPHER G. FAIRBURN

*Department of Psychiatry
University of Oxford*

G. TERENCE WILSON

*Graduate School of Applied and Professional
Psychology
Rutgers University, USA*

W. STEWART AGRAS

*Department of Psychiatry and Behaviour Sciences,
Stanford University, USA*

fore, internal calibration by appropriate comparison groups is an overriding necessity. The simple fact is that cognitive therapy has not as yet demonstrated that it is clearly superior to placebo in a sample that has also been shown to be medication-responsive.

That this design necessity cuts across diagnosis is indicated by Black *et al.* (1993), who assessed the cognitive therapy of panic versus fluvoxamine versus placebo. There have been many reports that cognitive therapy of panic is remarkably successful. However, in this trial (the only one that compares cognitive therapy to both medication and placebo) cognitive therapy was barely distinguishable from placebo, whereas fluvoxamine was markedly superior. This design is capable of cutting through much ambiguity and wishful thinking. Elkin *et al.* is a methodological standard that the field should adhere to.

- BLACK, D. W., WESNER, R., BOWERS, W., *et al.* (1993) A comparison of fluvoxamine, cognitive therapy, and placebo in the treatment of panic disorder. *Archives of General Psychiatry*, **50**, 44–50.

- ELKIN, I., SHEA, M., WATKINS, J., *et al.* (1992) National Institute of Mental Health treatment of depression collaborative treatment programme. *Archives of General Psychiatry*, **46**, 971–982.

DONALD F. KLEIN

*Columbia University
College of Physicians and Surgeons
Unit 22, 722 West 168th Street
New York 10032*

SIR: Scott, in her article on cognitive therapy (*BJP*, July 1994, **165**, 126–130), expresses surprise that the NIMH treatment study (Elkin *et al.*, 1992) indicates that CBT was barely distinguishable from placebo. She states: “It is hard to reconcile these findings with all the previous data (27 outcome studies) supporting the efficacy of CBT in major depression.”

However, 27 experimentally inadequate studies do not outweigh one well-designed one. Elkin *et al.* (1992) is the only study that made certain that the sample studied was medication-relevant, as shown by the superiority of medication to placebo. Having thus calibrated the sample, the insignificant difference of CBT from placebo achieves trenchant significance.

Studies that do not include both a placebo and a medication arm are irretrievably ambiguous. Unfortunately, our diagnostic rubrics allow for enormous heterogeneity and are only loosely linked to the prediction of therapeutic effect. There-

Childhood abuse and psychosis

SIR: In their study on childhood abuse in first-episode psychosis, Greenfield *et al.* (*BJP*, June 1994, **164**, 831–834) found that history of abuse was associated with significantly more dissociative symptoms. Unfortunately, the article does not indicate the extent to which the subjects were aware of the abuse aspect of the study at the time they completed the dissociation questionnaire. Council (1993) found that childhood trauma and dissociation were correlated only when the trauma survey preceded the measurement of dissociation, suggesting that the relationship between these variables may be an artefact of the context within which they are assessed. It should also be noted that the significant correlation between dissociation and trauma has been shown to disappear after statistically controlling for family pathology (Nash *et al.*, 1993).

I would also question the authors' conclusion that reliance on self-reports of abuse would more likely take the form of underreporting than over-reporting. Although there is evidence that this may occur in a non-clinical sample (Femina *et al.*, 1990), it is conceivable that an opposite pattern might emerge in a clinical sample, especially with the current emphasis by therapists and the media on the damaging effects of childhood sexual abuse. One must also consider the growing number of individuals who claim to have recently uncovered previously forgotten memories of abuse. Insofar as the reliability of these "recovered memories" is now being seriously questioned (e.g. Loftus, 1993), future research on this topic should regularly include memory continuity as an additional variable. Even if one assumes that recovered memories are reliable, they may be associated with a distinctly different pattern of symptoms. Moreover, as the controversy over recovered memories plays itself out, there is a need for an ongoing assessment of the frequency with which patients are reporting such memories.

COUNCIL, J. R. (1993) Context effects in personality research. *Current Directions in Psychological Science*, 2, 31–34.

FEMINA, D. D., YEAGER, C. A. & LEWIS, D. O. (1990) Child abuse: adolescent records vs. adult recall. *Child Abuse and Neglect*, 14, 227–231.

LOFTUS, E. B. (1993) The reality of repressed memories. *American Psychologist*, 48, 518–537.

NASH, M. R., HULSEY, T. L., SEXTON, M. C., *et al.* (1993) *Journal of Consulting and Clinical Psychology*, 61, 276–283.

RUSSELL A. POWELL

Department of Social Sciences
Grant MacEwan College
P.O. Box 1796
Edmonton, Alberta
Canada T5J 4S2

Community psychiatric nurse teams

SIR: Muijen *et al.* (*BJP*, August 1994, 165, 211–217), in discussing the results of their innovative and important clinical trial of standard v. intensive care by community psychiatric nurses (CPNs), concluded that "despite the low number of contacts in the generic [standard] group and high number in the CST [intensive] group, no differences in outcome were found". Although they interpreted these findings cautiously, some methodological considerations were omitted in their discussion of the possible explanations.

Randomisation, confounding, and type II error rate. Simple randomisation is a powerful method,

because it controls for both known and unknown confounding factors. This procedure, however, requires a large sample size, especially if one is trying to detect relatively small effects against potentially confounding background noise. The sample size used by Muijen *et al.* (82 at baseline) is rather small for this purpose, and consequently there remain some baseline differences between the two groups after randomisation. Therefore, it becomes vital to control adequately for confounding. The presence of some possible confounders, such as the higher proportion of CST patients living alone and being of Afro-Caribbean origin, are acknowledged. However, in spite of the already small sample size, the authors examine for possible confounding using the method of *post hoc* stratification, which results in very small subgroups. Such analyses are unlikely to be informative, break the simple randomisation, and should be avoided if a trial has limited statistical power (Pocock *et al.*, 1987a).

The lack of power in this study is compounded by the relatively high attrition rate (nearly 25%). Absence of "statistically significant differences" should, therefore, not be equated with truly negative results. Instead, the possibility of a type II error (i.e. a false negative result) should be considered. The issue of low power can be clarified by listing the appropriate confidence intervals: a statistically insignificant mean difference with a narrow confidence interval is an informative null finding, but not a statistically insignificant mean difference with a very wide one. Such data would enable the reader to evaluate whether the negative results are genuine or due to small sample size.

Analyses. The authors specify seven end points (considered of equal importance from the statistical point of view), and conduct separate *t*-tests of mean changes in each endpoint, and further separate analyses at each of the three measurement points. Such overuse of significance testing may make any possible results difficult to interpret. Furthermore, because of their larger standard error, analysis involving the mean change between baseline and post-treatment measurements are inferior to methods using the mean baseline measurement as a covariate in a linear model for treatment comparisons of post-treatment means (Frison & Pocock, 1992). The method used will also fail to pick up the fact that, for example, there were small but consistent differences between the two groups on patient satisfaction scores (for a discussion of suitable methods see Pocock *et al.*, 1987b; Frison & Pocock, 1992). It needs to be pointed out that each test is conducted on slightly different samples, because of